

OIPE

RAW SEQUENCE LISTING DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

Input Set : A:\FimH Peptides.ST25.txt Output Set: N:\CRF3\07232001\I900575.raw

```
ENTERED
      3 <110> APPLICANT: Langermann, Solomon
              Revel, Andrew
      5
              Auguste, Christine
      6
              Burlein, Jeanne
      8 <120> TITLE OF INVENTION: FimH Adhesin Proteins and Methods of Use
     10 <130> FILE REFERENCE: 469201-549
C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/900,575
C--> 12 <141> CURRENT FILING DATE: 2001-07-06
     12 <150> PRIOR APPLICATION NUMBER: US/60/216,750
     13 <151> PRIOR FILING DATE: 2000-07-07
     15 <160> NUMBER OF SEQ ID NOS: 64
     17 <170> SOFTWARE: PatentIn version 3.0
     19 <210> SEQ ID NO: 1
     20 <211> LENGTH: 837
     21 <212> TYPE: DNA
     22 <213> ORGANISM: E. coli
     24 <400> SEQUENCE: 1
     25 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt
                                                                               60
     27 tatgtaaacc ttgcgcccgt cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg
                                                                              120
     29 caaatetttt gecataaega ttateeggaa accattaeag actatgteae actgeaaega
                                                                              180
     31 ggctcggctt atggcggcgt gttatctaat ttttccggga tcgtaaaata tagtggcagt
                                                                              240
     33 agctateett teeetaeeae eagegaaaeg eegegegttg tttataatte gagaaeggat
                                                                              300
     35 aageegtgge eggtggeget ttatttgaeg eetgtgagea gtgegggggg agtggegatt
                                                                              360
     37 aaagcagget cattaattge egtgettatt ttgegacaga ceaacaacta taacagegat
                                                                              420
     39 ggtttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc
                                                                              480
    41 tgcgatgett ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgccg
                                                                              540
    43 attectetta cegittatig igegaaaage caaaaeetgg ggiattaeet eteeggeaca
                                                                              600
     45 acegeaggtg egggeaacte gatttteace aatacegegt egtttteace egegeaggge
                                                                              660
     47 gteggegtae agttggegeg caaeggtaeg gttatteeag egaataaeae ggtategtta
                                                                              720
    49 ggagcagtag ggacttegge ggtgagtetg ggattaaegg caaattaege aegtaeegga
                                                                              780
    51 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa
                                                                              837
    54 <210> SEQ ID NO: 2
    55 <211> LENGTH: 837
    56 <212> TYPE: DNA
    57 <213> ORGANISM: E. coli
    59 <400> SEQUENCE: 2
    60 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt
                                                                               60
    62 tatgtaaacc ttgcgcctgc cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg
                                                                              120
    64 caaatetttt gecataaega ttaeeeggaa aeeattaeag aetatgteae aetgeaaega
                                                                              180
    66 ggttcggctt atggcggcgt gttatctagt ttttccgggga tcgtaaaata taatggcagt
                                                                              240
    68 agctatectt teectactae cagegaaaeg eegegggttg tttataatte gagaaeggat
                                                                              300
    70 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcggggg agtggcgatt
                                                                              360
    72 aaagcagget cattaattge egtgettatt ttgegacaga ccaacaacta taacagegat
                                                                              420
    74 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc
                                                                              480
    76 tgcgatgctt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgccg
                                                                              540
    78 attectetta cegittatty tycyaaaage caaaacetyy yytattaeet eteegyeaca
                                                                              600
    80 accgcagatg cgggcaactc gattttcacc aataccgcgt cgttttcacc cgcgcagggc
```

660

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,575

DATE: 07/23/2001 TIME: 13:37:27

82	gtcggcgtac agttgacgcg	caacggtacg	attattccag	cgaataacac	ggtatcgtta	720
84	ggagcagtag ggacttcggc	ggtaagtctg	ggattaacgg	caaattacqc	acqtaccqqa	780
86	gggcaggtga ctgcagggaa	tgtgcaatcg	attattggcg	tgacttttgt	ttatcaa	837
89	<210> SEQ ID NO: 3			-		
	<211> LENGTH: 837					
	<212> TYPE: DNA					
	<213> ORGANISM: E. CO	li				
	<400> SEQUENCE: 3					*
95	ttcgcctgta aaaccgccaa	tggtacagct	atccctattg	gcggtggcag	cqctaatqtt	60
97	tatgtaaacc ttgcgcctgc	cgtgaatgtg	gggcaaaacc	tggtcgtaga	tetttegacg	120
99	caaatctttt gccataacga	ttatccggaa	accattacag	actatqtcac	actgcaacga	180
10.	l ggctcggctt atggcggcg	t gttatctaat	ttttccggga	a ccqtaaaata	tagtggcagt	240
103	3 agctatccat ttccgacca	c cagegaaacg	ccgcgggtto	, tttataatto	gagaacggat	300
105	aagccgtggc cggtggcgc	t ttatttgacg	r cctgtgagca	atacaaacaa	ggtggcgatt	360
107	' aaagctggct cattaattg	c cgtgcttatt	ttgcgacaga	ccaacaacta	taacagcgat	420
103) gatttccagt ttgtgtgga	a tatttacgcc	: aataatgatg	r taataatac	tactggcggc	480
111	. tgcgatgttt ctgctcgtg	a tgtcaccgtt	actctgccgg	actaccctqq	ttcagtgcca	540
113	attoctotta cogtttatt	g tgcgaaaagc	caaaacctgg	ggtattacct	ctccggcaca	600
115	accgcagatg cgggcaact	c gattttcacc	: aataccgcgt	cgttttcacc	agcgcagggc	660
117	gtcggcgttc agttgacgc	g caacggtacg	attattccca	cgaataacac	ggtatcgtta	720
119	ggagcagtac ggacttcgg	c ggtaagtctg	ggattaacqq	caaattacqc	acqtaccqqa -	780
121	. gggcaggtga ctgcagggaa	a tgtgcaatcg	attattggcg	tgacttttgt	ttatcaa	837
124	<210> SEQ ID NO: 4			-		
	<211> LENGTH: 840					
	<212> TYPE: DNA					
	<213> ORGANISM: E. co	oli				
	<400> SEQUENCE: 4					
130	ttcgcctgta aaaccgccaa	a tggtaccgca	atccctattg	gcggtggcag	cgccaatgtt	60
132	tatgtaaacc ttgcgcctgo	cgtgaatgtg	gggcaaaacc	tggtcgtaga	tetttegacg	120
134	caaatctttt gccataacga	ı ttacccagaa	accattacag	actatqtcac	actgcaacga	180
136	ggtgcggctt atggcggcgt	gttatctagt	ttttccggga	ccgtaaaata	taatggcagt	240
138	agctatcctt tccctactac	cagcgaaacg	ccgcgggttg	tttataattc	gagaacggat	300
140	aagccgtggc cggtggcgct	: ttatttgacg	ccggtgagca	atacaaaaaa	agtggcgatt	360
142	aaagctggct cattaattgo	: cgtgcttatt	ttgcgacaga	ccaacaacta	taacagcgat	420
144	gatttccagt ttgtgtggaa	ı tatttacgcc	aataatgatg	tggtggtgcc	cactggcggc	480
146	tgcgatgttt ctgctcgtga	tgtcaccgtt	actctgccgg	actaccctqq	ttcagtgccg	540
148	attectetta cegtttattg	tgcgaaaagc	caaaacctgg	ggtattacct	ctccggcaca	600
150	accgcagatg cgggcaacto	gattttcacc	aataccgcgt	cqttttcacc	cgcgcagggc	660
152	gtcggcgtac agttgacgcg	caacqqtacq	attattccag	cgaataacac	ggtatcgtta	720
154	ggagcagtag ggacttcggc	ggtaagtctg	ggattaacgg	caaattacgc	acqtaccqqa	780
120	gggcaggtga ctgcagggaa	tgtgcaatcg	attattggcg	tgacttttgt	ttatcaataa	840
159	<210> SEQ ID NO: 5					
	<211> LENGTH: 840					
	<212> TYPE: DNA					
	<213> ORGANISM: E. co	li				
	<400> SEQUENCE: 5					
165	ttcgcctgta aaaccgccaa	tggtaccgct	attcctattg	gcggtggcag	cgctaatgtt	60
T6/	tatgtaaacc ttgcgcctgc	cgtgaatgtg	gggcaaaacc	tggtcgtaga	tctttcgacg	120
169	caaatctttt gccataacga	ttatccggaa	accattacag	actatgtcac	actgcaacga	180

RAW SEQUENCE LISTING

DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

171 ggctcggctt atggggggt attacks	
171 ggctcggctt atggcggcgt gttatctaat ttttccggga ccgtaaaata tagtggcagt	240
	300
	360
	420
	480
	540
	600
	660
	720
	780
JJJ~~JJ~J~J~~J~~~~~~~~~~~~~~~~~~~~~~~~	840
• • • • • • • • • • • • • • • • • •	040
195 <211> LENGTH: 837	
196 <212> TYPE: DNA	
197 <213> ORGANISM: E. coli	
199 <400> SEQUENCE: 6	
200 ttcgcctgta aaaccgccaa tggcaccgct atccctattg gcggtggcag cgccaatgtt	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
January and an ended the control of	660
	720
- JJJ~~JJ~94 CC4CQ444Q	780
> (210) OHO ID NO: /	837
230 <211> LENGTH: 837	
231 <212> TYPE: DNA	
232 <213> ORGANISM: E. coli	
234 <400> SEQUENCE: 7	
235 ttcgcctgta aaaccgccaa tggtaccgct atccctattg gcggtggcag cgccaatgtt	
	60
	120
	180
	240
	300
	360
	420
	480
	540
	600
	660
	720
261 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa	780
The second secon	837

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/900,575

DATE: 07/23/2001 TIME: 13:37:27

264 <210> SEQ ID NO: 8	
265 <211> LENGTH: 837	
266 <212> TYPE: DNA	
267 <213> ORGANISM: E. coli	
269 <400> SEQUENCE: 8	
270 tttgcctgta aaaccgccaa tggcaccgct atccctattg gcggtggcag cgccaatgtt	C 0
272 tatgladact tyycycccyc cyfgaafgfg gggcaaaacc fggfggfgga fottfgga	60
2/4 cadacetttt geedlaaega ttateeggaa accattacag actatoteag actagas	120
270 ggoodggood alggoggogt attatotaat tittooggaa cogtaaaata tagtagaata	180
270 ayotatodat ticoqactae caqeqaaacq coqeqqatta tttataatta qaqaaaaat	240
200 dayccytyyc cyglygcycl tlattigacg cctgtgagca gtgcgggtgg cgtgggcy	300
202 adayctyyct cattaattqc cqtqcttatt ttqcqacaga ccaacaacta taacaacaa	360
284 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc tactggcggc	420
286 tgcgatgttt ctgctcatga tgtcaccgtt actctgccgg actaccctgg ttcagtgcca	480
288 attectetta cegittatta taggaaaaage caaaacetga gatattacet eteeggeaca	540
290 accgcagatg cgggcaactc gattttcacc aataccgcgt cgttttcacc agcgcagggc	600
292 gtcggcgtac agttgacgcg caacggtacg attattccag cgaataacac ggtatcgtta	660
294 ggagcagtag ggacttcggc ggtgagtctg ggattaacgg caaattacgc acgtaccgga	720
296 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa	780
299 <210> SEQ ID NO: 9	837
300 <211> LENGTH: 837	
301 <212> TYPE: DNA	
302 <213> ORGANISM: E. coli	
304 <400> SEQUENCE: 9	
305 ttcgcctgta aaaccgccaa tggtaccgca atccctattg gcggtggcag cgccaatgtt	
307 tatgtaaacc ttgcgcctgc cgtgaatgtg gggcaaaacc tggtcgtaga tctttcgacg	60
309 canatotttt gocataacga ttaccagaa accattacag actatgtcac actgcaacga	120
311 ggttcggctt atggcggcgt gttatctagt ttttccggga ccgtaaaata taatggcagt	180
313 agctatectt tecetactae cagegaaacg cegegggttg tttataatte gagaacggat	240
315 aagccgtggc cggtggcgct ttatttgacg ccggtgagca gtgcggggg agtggcgatt	300
317 aaagctggct cattaattgc cgtgcttatt ttgcgacaga ccaacaacta taacagcgat	360
319 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc	420
321 tgtgatgett etgetegtga tgteacegtt actttgeegg actaceetgg tteagtgeeg	480
323 attectetta cegittatig tgegaaaage caaaacetgg ggtattacet ateeggeaca	540
325 accgcagate egggcaactc gattttcacc aataccgcgt cgttttcacc egcgcagggc	600
327 gteggegtac agttgaegeg caaeggtaeg attatteeag egaataacae ggtategtta	660
329 ggagcagtag ggacttcggc ggtaagtctg ggattaacgg caaattacgc acgtaccgga	720
331 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa	780
334 <210> SEQ ID NO: 10	837
335 <211> LENGTH: 840	
336 <212> TYPE: DNA	
337 <213> ORGANISM: E. coli	
339 <400> SEQUENCE: 10	
340 ttcgcctgta aaaccgccaa tggcagggt atggatat	
340 ttegeetgta aaacegecaa tggeaceget atecetattg geggtggeag egeeaatgtt 342 tatgtaaace ttggegeege egtgaatgtg gggcaaaace tggtegtgga tetttegaeg	60
344 caaatettt gecataaega ttaeceggaa accattacag attatgteac actgeaaega	120
346 ggctcggctt atggcggcgt gttatotaat ++++	180
346 ggctcggctt atggcggcgt gttatctaat ttttccggga ccgtaaaata tagtggcagt 348 agctatccat ttccgaccac cagtgaaaag caggaggtata	240
348 agctatccat ttccgaccac cagtgaaacg ccgcgggttg tttataattc gagaacggat 350 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcgggcgg ggtggtgatt	300
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	360

RAW SEQUENCE LISTING

DATE: 07/23/2001 PATENT APPLICATION: US/09/900,575 TIME: 13:37:27

_ (= : : : : : : : : : : : : : : : : : :	
352 aaagetgget cattaattge egtgettatt ttgegacaga ceaacaacta taacagegat	
354 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc	420
	480
	540
360 accgcagatg cgggcaactc gattttcacc aataccgcgt cgttttcacc tgcacagggc	600
	660
364 ggagcagtag ggacttcggc ggtaagtctg ggattaacgg caaattacgc acgtaccgga 366 gggcaggtga ctgcagggaa tgtgcaatag attattc	720
	780
	840
370 <211> LENGTH: 837	
371 <212> TYPE: DNA	
372 <213> ORGANISM: E. coli	
374 <400> SEQUENCE: 11	
375 ttegeetgta aaacegeeaa tggtacegea atcectattg geggtggeag egeeaatgtt	
	60
	120
	180
383 agetateett teettaetae cagegaaaeg cegegggttg tttataatte gagaaeggat	240
385 aagccgtggc cggtggcgct ttatttgacg ccggtgagca gtgcggggg agtggcgatt	300
387 aaagctggct cattaattgc cgtgcttatt ttgcgacaga ccaacaacta taacagcgat	360
389 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc cactggcggc 391 tgcgatgttt ctgctcgtga tgtcaccgtt agtataaa	420
391 tgcgatgttt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgccg	480
	540
	600
	660
	720
401 gggcaggtga ctgcagggaa tgtgcaatcg attattggcg tgacttttgt ttatcaa 404 <210> SEO ID NO: 12	780
404 <210> SEQ ID NO: 12	837
405 <211> LENGTH: 840	
406 <212> TYPE: DNA	
407 <213> ORGANISM: E. coli	
409 <400> SEQUENCE: 12	
410 ttcgcctgta aaaccgccaa tggtagggt abandus	
412 tatgtaaacc ttgcgcccgt cgtgaatgtg gggcaaaacc tggtcgtgga tctttcgacg	60
414 caaatetttt gecataaega ttateeggaa aceattaeag actatgteae actgeaaega 416 ggeteggett atggeggegt gttatetaat ttate	120
416 ggctcggctt atggcggcgt gttatctaat ttttccggga ccgtaaaata tagtggcagt 418 agctatccat ttcctaccac cagggaaaaa	180
	240
420 aagccgtggc cggtggcgct ttatttgacg cctgtgagca gtgcgggcgg ggtggcgatt	300
422 aaagetgget cattaattge egtgettatt ttgegacaga ecaacaacta taacagegat 424 gattteeagt ttgtgtggaa tatttaggag aataatta	360
424 gatttccagt ttgtgtggaa tatttacgcc aataatgatg tggtggtgcc tactggcggc	420
426 tgcgatgttt ctgctcgtga tgtcaccgtt actctgccgg actaccctgg ttcagtgcca 428 attcctctta ccgtttattg tgcgaaaagg canacacttg	480
428 attectetta cogtttatta tacquanaga anno actaccetga tteagtgeea	540
430 accgcagatg cgggcaactc gattttcacc antagagett ggtattacct ctccggcaca	600
432 gtcggcgtac agttgacgcg caacagtaca attaccgcgt cgttttcacc tgcacagggc	660
434 ggagcagtag ggacttcggc ggtgagtctg grattattccag cgaataacac ggtatcgtta	720
436 gggcaggtga ctgcagggaa tgtgcaatog attatacgg caaattatgc acgtaccgga	780
439 <210> SEQ ID NO: 13	840
440 <211> LENGTH: 837	
441 <212> TYPE: DNA	

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/900,575

DATE: 07/23/2001
TIME: 13:37:28

Input Set : A:\FimH Peptides.ST25.txt
Output Set: N:\CRF3\07232001\1900575.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date L:2195 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:46 L:2453 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:47 L:2489 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:48 L:2531 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:49 L:2569 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:50 L:2654 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:52 L:2732 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:55 L:2796 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:56 L:2808 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:57 L:2820 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:58 L:2832 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:59 L:2844 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:60 L:2856 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:61 L:2868 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:62 L:2880 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:63 L:2892 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:64